

CLAIMS

1. A machine-based method comprising
receiving historical multi-dimensional data representing multiple source variables
to be used as an input to a predictive model of a commercial system,
applying transformations to the source variables that are selected to increase
predictive power, and
applying transformations to the data that are selected based on the strength of
measurement represented by a variable.
2. The method of claim 1 in which the strength of measurement comprises at least
one of nominal, ordinal, and interval.
3. The method of claim 1 in which the strength of a measurement is represented in
stored metadata associated with the data.
4. The method of claim 1 also including
displaying to a user a representation of a response function of a target variable against
untransformed, transformed, and target variables associated with the data.
5. The method of claim 1 also including
persistently storing both the source variables and related transformed versions of the
source variables.
6. A machine-based method comprising
receiving historical multi-dimensional data representing multiple source variables
to be used as an input to a predictive model of a commercial system,
adjusting unstable values of the variables to reduce inaccurate associations between
predictor variables and target variables.
7. The method of claim 6 in which the adjustment of the unstable values comprises
Bayesian renormalization.
8. A machine-based method comprising
in connection with a project in which a user generates a predictive model based on
historical data about a system being modeled, automatically imputing missing values for
continuous variables associated with the data.

9. The method of claim 8 in which the user is enabled to invoke the automatic imputing as part of a user interface feature that displays information about variables for which values are missing.
10. The method of claim 9 in which the automatic imputing is invoked based on the variable or type of variable.
11. The method of claim 9 in which the variables for which missing values are imputed may be used in the model or may be transformed for use in the model.